



(11) Publication number: **0 431 960 A2**

(12)

EUROPEAN PATENT APPLICATION

(21) Application number: **90313311.4**

(51) Int. Cl.⁵: **H04N 1/40**

(22) Date of filing: **07.12.90**

(30) Priority: **08.12.89 US 448193**

(43) Date of publication of application:
12.06.91 Bulletin 91/24

(84) Designated Contracting States:
DE FR GB

(71) Applicant: **XEROX CORPORATION**
Xerox Square - 020
Rochester New York 14644 (US)

(72) Inventor: **Bloomberg, Dan S.**
1013 Paradise Way
Palo Alto, California, 94306 (US)
Inventor: **Crocca, William T.**
114 Heather Way
Rochester, New York, 14625 (US)

(74) Representative: **Hill, Cecilia Ann et al**
Rank Xerox Patent Department Albion House,
55 New Oxford Street
London WC1A 1BS (GB)

(54) Identification, characterization, and segmentation of halftone or striped regions of binary images.

(57) A method for creating a mask for separating halftone regions in a binary image from other regions comprises: constructing a seed image that includes pixels only in halftone regions and at least one pixel in every halftone region (67); constructing a clipping mask that covers in a connected manner all ON pixels in halftone regions (70); and filling the seed while clipping to the mask (72). Thresholded reductions and morphological operations are preferred.

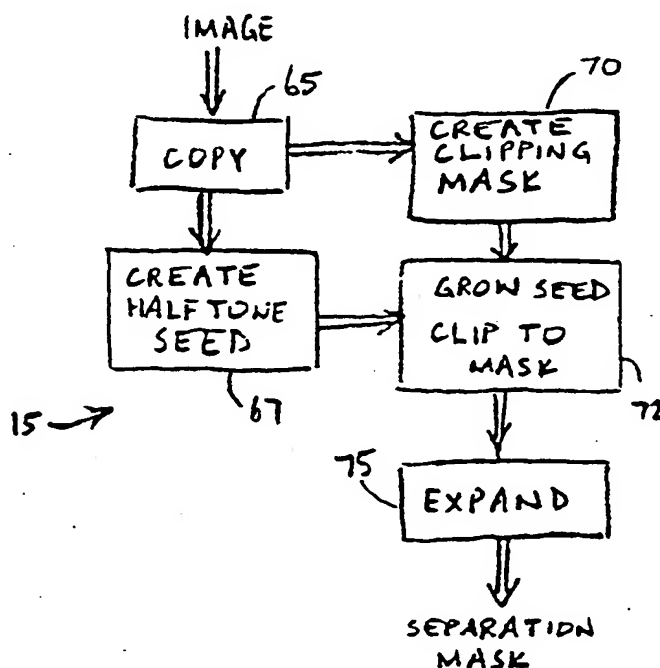


FIG. 9A

EP 0 431 960 A2

IDENTIFICATION, CHARACTERIZATION, AND SEGMENTATION OF HALFTONE OR STIPPLED REGIONS OF BINARY IMAGES

The invention relates generally to image processing, and more specifically to a technique for discriminating between finely textured regions and other regions.

There are several applications where it is important to determine quickly whether an image contains regions of fine texture such as halftones and stipples.

For example, problems can arise with printers whose output is at a different resolution from the input scanner, since techniques for converting the resolution are sensitive to texture. A technique that works well on text will generally do poorly on halftones or stipples, and vice versa. If the halftone regions are identified, and if the frequency and screen angle are known, then appropriate techniques can be selected for different portions of the image, and the halftone regions can be resolution converted with an acceptably low level of aliasing artifacts.

Problems also arise for scanners. When scanning a halftoned binary image, beating can occur between the repeat frequency in the image and the size of an integral number of scanned pixels. The result is aliasing, in which a low frequency Moiré pattern will be observed in the scanned image. In order to prevent this, it is desirable to use a gray scale scanner, and remove halftoning prior to thresholding. Techniques for removing halftone patterns work best if the halftone frequency and screen angle are known.

It is sometimes necessary for gross segmentation to occur prior to using some segmentation software. Some aspects of segmentation can be accomplished by building a connected component representation of the binary image, such as a line adjacency graph (LAG), and then processing that data structure. However, if one tried to build a LAG from a finely textured screen of appreciable size (say 2 inches square or larger), the storage requirements and computational time might well be excessive. Similarly, it is necessary for segmentation to occur prior to using recognition software. If a halftoned region of a binary image were sent to OCR or graphics vectorization software, it could break the program.

Depending on the threshold setting and resolution of the input scanners, and on the quality of the output printer and number of copy generations, a printed version of the binary image of a regular textured region may show little or none of the details of the texturing of the original binary image. Because of scanning and printing operations, the contrast at the 1.5 line pairs per mm size is often greatly increased, with dark textures becoming solid black, and light textures becoming much lighter or even white (perhaps with some random dot noise). Thus, in many situations, it cannot be assumed that evidence of the origi-

nal texturing will remain in the binary image under analysis.

It is an object of the present invention to enable a representation of the halftone regions of a binary image to be created, and facilitate efficient and effective separation of textured regions in the image from non-textured regions.

Techniques in accordance with the invention, described below, utilize transformations such as thresholded reductions and morphological operations, which will be defined and discussed in detail below. A thresholded reduction entails mapping a rectangular array of pixels onto a single pixel, whose value depends on the number of ON pixels in the rectangular array and a threshold level. Morphological operations use a pixel pattern called a structuring element (SE) to erode, dilate open, or close an image.

A preferred technique in accordance with the invention for determining whether there exist halftone regions includes dividing the image into a number of sub-regions or tiles, extracting and counting pixel transitions in each tile, normalizing the counts to the area, and comparing the values to a threshold value that distinguishes halftone regions from non-halftone regions (such as those containing text and line graphics).

A preferred technique in accordance with the invention for determining the screen size and angle preferably includes selecting the tile having the largest number of pixel transitions, and filtering the tile, such as by eroding it with a number of hit-miss SE's that act as narrow bandpass filters. The SE's are each characterized by a linear dimension and angle. The SE giving the most ON pixels after erosion of the tile serves to characterize the screen size and angle.

Constructing the separation mask preferably includes constructing a seed image, constructing a clipping mask, and filling the seed while clipping to the mask. The seed must contain pixels only in the halftone regions and must contain at least one pixel in every halftone region. The clipping mask must cover all ON pixels in the halftone regions. It may also cover non-halftone regions, but no part of the clipping mask that covers a non-halftone region should touch a part that covers a halftone region.

The seed may be constructed by eroding the entire image with the SE that was found to give the best match in connection with determining the screen size and angle as mentioned above, performing one or more thresholded reductions, and performing an open operation with a 3×3 or 4×4 solid SE to eliminate pixels outside the halftone regions. More specifically, in accordance with the invention, the step of constructing a seed comprises the steps of:

providing a set of spatial filters ;
 selecting the spatial filter that best matches the
 halftone characteristics of at least a portion of the
 image ;
 eroding the original binary image with the selec-
 ted spatial filter ;
 subjecting the image, thus filtered, to set of oper-
 ations operation that eliminates OFF pixels that
 are near ON pixels ; and
 subjecting the resulting image to an open oper-
 ation.

The clipping mask may be constructed by con-
 verting the halftone regions to solid ON pixels. This
 is preferably accomplished by subjecting the entire
 image to a transformation that tends to eliminate OFF
 pixels located next to or near ON pixels in a manner
 that the halftone regions become substantially
 entirely filled, while the non-halftone regions (such as
 those containing line graphics and text) become dar-
 kened but not completely filled. By way of example,
 such a transformation may comprise one or more
 thresholded reductions with general darkening (low
 threshold level), possibly followed by a close oper-
 ation. Accordingly, the step of constructing a clipping
 mask comprises at least one thresholded reduction to
 solidify textured areas.

Filling the seed while clipping to the mask may be
 accomplished by an iterative procedure wherein the
 seed (or current iteration thereof) is dilated with a 3x3
 solid SE and ANDed with the clipping mask until the
 result does not change.

The present invention accordingly provides a
 method for constructing a separation mask for
 halftone regions in an original binary image, compris-
 ing the steps of :

processing the original image to construct a seed
 image, the seed image having the property that it
 contains pixels only in the halftone regions and
 that it contains at least one pixel in every halftone
 region ;
 processing the original image to form a clipping
 mask, the clipping mask having the property that
 it covers all ON pixels in the halftone regions and
 that any part of the clipping mask that covers a
 non-halftone region does not touch a part of the
 clipping mask that covers a halftone region ; and
 processing the seed image and the clipping mask
 so as to grow the seed and clip to the clipping
 mask ;
 whereby the seed, thus grown and clipped pro-
 vides a representation of the separation mask.
 The step of constructing a seed may comprise the
 steps of :
 providing a set of spatial filters ;
 selecting the spatial filter that best matches the
 halftone characteristics ;
 eroding the original binary image with the selec-
 ted spatial filter ;

subjecting the image, thus filtered, to at least one
 thresholded reduction ; and
 subjecting the resulting image to an open oper-
 ation.

By way of example only, embodiments of the pre-
 sent invention will now be described with reference to
 the accompanying drawings, in which :

Fig. 1A is a block diagram of an image scanning
 and processing system ;

Fig. 1B is a high level flow diagram showing the
 method of identifying, characterizing, and
 separating finely textured regions in an image ;
 Figs. 2A-B are flow diagrams illustrating the steps
 in determining whether there are halftones ;

Fig. 3 is a flow diagram illustrating the characteri-
 zation of the size and angle of the halftone
 screen ;

Figs. 4A-D, 5A-C, 6A-D, 7A-B, and 8A-D show the
 hit-miss SE's that are used as narrow bandpass
 filters to determine the halftone characteristics ;
 Fig. 9A-D are flow diagrams illustrating the crea-
 tion of a separation mask from a seed and a clip-
 ping mask ;

Fig. 10 is a flow diagram illustrating the use of the
 mask to obtain the separation of the textured and
 non-textured regions ;

Fig. 11 is a flow diagram illustrating an alternative
 to expanding the mask ;

Fig. 12 is a flow diagram illustrating an alternative
 way to create the seed ;

Figs. 13A-B are flow diagrams illustrating alterna-
 tives to the thresholded reductions used to create
 the seed and the clipping mask ;

Fig. 14 shows a representative binary image ;

Fig. 15A is a plot of filter response for a set of spa-
 tial filters ;

Figs. 15B-D show the results at different stages
 of creating of the seed ;

Figs. 16A-B show the results at different stages
 of creating the clipping mask ;

Figs. 17A-B show the results of combining the
 seed and clipping mask to provide a halftone
 mask ;

Figs. 18A-B show the halftone separation and the
 text separation ; and

Fig. 19 is a block diagram of special purpose
 hardware for performing image reductions and
 expansions.

Definitions and Terminology

The present discussion deals with binary images.
 In this context, the term "image" refers to a represen-
 tation of a two-dimensional data structure composed
 of pixels. A binary image is an image where a given
 pixel is either "ON" or "OFF." Binary images are mani-
 pulated according to a number of operations wherein
 one or more source images are mapped onto a des-

tionation image. The results of such operations are generally referred to as images. The image that is the starting point for processing will sometimes be referred to as the original image.

Pixels are defined to be ON if they are black and OFF if they are white. It should be noted that the designation of black as ON and white as OFF reflects the fact that most documents of interest have a black foreground and a white background. While the techniques to be described could be applied to negative images as well, the discussion will be in terms of black on white.

A "solid region" of an image refers to a region extending many pixels in both dimensions within which substantially all the pixels are ON.

A "textured region" of an image refers to a region that contains a relatively fine-grained pattern. Examples of textured regions are halftoned or stippled regions.

"Text" refers to portions of a document or image containing letters, numbers, or other symbols including non-alphabetic linguistic characters.

"Line graphics" refers to portions of a document or image composed of graphs, figures, or drawings other than text, generally composed of horizontal, vertical, and skewed lines having substantial run length as compared to text. Graphics could range from, for example, horizontal and vertical lines in an organization chart to more complicated horizontal, vertical, and skewed lines in engineering drawings.

A "mask" refers to an image, normally derived from an original image, that contains substantially solid regions of ON pixels corresponding to regions of interest in the original image. The mask may also contain regions of ON pixels that don't correspond to regions of interest.

AND, OR, and XOR are logical operations carried out between two images on a pixel-by-pixel basis.

NOT is a logical operation carried out on a single image on a pixel-by-pixel basis.

"Expansion" is a scale operation characterized by a SCALE factor N, wherein each pixel in a source image becomes an NxN square of pixels, all having the same value as the original pixel.

"Reduction" is a scale operation characterized by a SCALE factor N and a threshold LEVEL M. Reduction with SCALE = N entails dividing the source image into NxN squares of pixels, mapping each such square in the source image to a single pixel on the destination image. The value for the pixel in the destination image is determined by the threshold LEVEL M, which is a number between 1 and N². If the number of ON pixels in the pixel square is greater or equal to M, the destination pixel is ON, otherwise it is OFF.

"Subsampling" is an operation wherein the source image is subdivided into smaller (typically square) elements, and each element in the source image is mapped to a smaller element in the desti-

nation image. The pixel values for each destination image element are defined by a selected subset of the pixels in the source image element. Typically, subsampling entails mapping to single pixels, with the destination pixel value being the same as a selected pixel from the source image element. The selection may be predetermined (e.g. upper left pixel) or random.

A "4-connected region" is a set of ON pixels wherein each pixel in the set is laterally or vertically adjacent to at least one other pixel in the set.

An "8-connected region" is a set of ON pixels wherein each pixel in the set is laterally, vertically, or diagonally adjacent to at least one other pixel in the set.

A number of morphological operations map a source image onto an equally sized destination image according to a rule defined by a pixel pattern called a structuring element (SE). The SE is defined by a center location and a number of pixel locations, each having a defined value (ON or OFF). Other pixel positions, referred to as "don't care," are ignored. The pixels defining the SE do not have to be adjacent each other. The center location need not be at the geometrical center of the pattern; indeed it need not even be inside the pattern.

A "solid" SE refers to an SE having a periphery within which all pixels are ON. For example, a solid 2 x 2 SE is a 2 x 2 square of ON pixels. A solid SE need not be rectangular.

A "hit-miss" SE refers to an SE that specifies at least one ON pixel and at least one OFF pixel.

"Erosion" is a morphological operation wherein a given pixel in the destination image is turned ON if and only if the result of superimposing the SE center on the corresponding pixel location in the source image results in a match between all ON and OFF pixels in the SE and the underlying pixels in the source image.

"Dilation" is a morphological operation wherein a given pixel in the source image being ON causes the SE to be written into the destination image with the SE center at the corresponding location in the destination image. The SE's used for dilation typically have no OFF pixels.

"Opening" is a morphological operation that consists of an erosion followed by a dilation. The result is to replicate the SE in the destination image for each match in the source image.

"Closing" is a morphological operation consisting of a dilation followed by an erosion.

The various operations defined above are sometimes referred to in noun, adjective, and verb forms. For example, references to dilation (noun form) may be in terms of dilating the image or the image being dilated (verb forms) or the image being subjected to a dilation operation (adjective form). No difference in meaning is intended.

System Overview

Fig. 1A is a block diagram of an image analysis system 1. The basic operation of system 1 is to extract or eliminate certain characteristic portions of a document 2. To this end, the system includes a scanner 3 which digitizes the document on a pixel basis, and provides a resultant data structure, typically referred to as an image. Depending on the application, the scanner may provide a binary image (a single bit per pixel) or a gray scale image (a plurality of bits per pixel). The image contains the raw content of the document, to the precision of the resolution of the scanner. The image may be sent to a memory 4 or stored as a file in a file storage unit 5, which may be a disk or other mass storage device.

A processor 6 controls the data flow and performs the image processing. Processor 6 may be a general purpose computer, a special purpose computer optimized for image processing operations, or a combination of a general purpose computer and auxiliary special purpose hardware. If a file storage unit is used, the image is transferred to memory 4 prior to processing. Memory 4 may also be used to store intermediate data structures and possibly a final processed data structure.

The result of the image processing can be a derived image, numerical data (such as coordinates of salient features of the image) or a combination. This information may be communicated to application specific hardware 8, which may be a printer or display, or may be written back to file storage unit 5.

Overview of the Separation of Halftone Regions

An original binary image may include regions of fine texturing such as stippling or halftoning. Such regions will be referred to as halftone regions. The image may also contain solid white regions, solid black regions, and regions of text and line graphics. Such regions will sometimes be referred to as non-halftone regions or other regions.

Fig. 1B is a high level flow diagram showing the major steps in a procedure for identifying, characterizing, and separating halftone or stippled regions in an original binary image. In brief, the technique entails subjecting the original binary image to a series of logical, scale, and morphological operations. The original binary image is first processed to determine whether there are halftone regions (step 10), and operation is ceased if there are not. Assuming there are, the image is subjected to a filtering procedure to determine the characteristic size and angle of the halftone screen (step 12). The image is then processed to provide a separation mask that covers the halftone regions and no other regions (step 15). The mask and the original binary image are then combined to provide the separations of the halftone and non-halftone regions, and

coordinates of the halftone regions are identified (step 20).

Detailed Discussion of One Embodiment of the Invention

Fig. 2A is an expanded flow diagram illustrating the steps within step 10 (determining whether there are halftone regions). The image is first divided into a number of sub-regions or tiles (step 22). Each tile must contain many pixels and should be on the order of the smallest expected halftone region. Specifically, the image may be divided into tiles, each of a size of about 100x100 pixels to 250x250 pixels. For each tile, a number is generated that represents the number of pixel transitions in that tile (step 25), and the tile with the maximum number of pixel transitions is selected (step 27).

An excellent way to discriminate between halftone and non-halftone regions is to count the number of transition pixels within a region, and to divide that number by the area of the region. Normalizing to the area of the region takes into account two factors that differentiate non-halftone from halftone regions: texture size for pixel transitions and fill fraction of the textured pattern. On both counts, halftone regions give results that are significantly larger than non-halftone regions, even those containing text, typically allowing an easy determination. A preferred measure for discrimination is to use the number of horizontal pixel transitions (ONOFF or OFFON along the scanline), and to take the ratio of such pixel transitions to the number of 16-bit words in the tile. A region of text typically has a value for this ratio of less than 2.0 and a halftone region is typically greater than 2.5.

Accordingly, once the tile with the maximum number has been selected, that number is normalized and a determination is made whether the normalized pixel transition number exceeds a threshold (step 30). If not, it is assumed there are no halftone regions, and processing stops. If the tile with the maximum number of pixel transitions does exceed the threshold, it is assumed that there are halftone regions and processing continues.

Fig. 2B is a flow diagram illustrating a particular technique for carrying out the counting of horizontal pixel transitions (step 25). The operations in Fig. 2B are carried out for each tile resulting from step 22. The tile is copied (step 35) and one copy reserved. The other copy is eroded (step 40) with a 1 x 3 horizontal SE 37 having its center at the center position and the result is XORed with the reserved copy (step 42). These operations have the effect of mapping pixel transitions to edge pixels, that is, pixels which are adjacent to a pixel of the opposite color. This technique can slightly undercount pixel transitions, but that is not a significant problem. Alternatively, vertical transitions could be counted by using a 3 x 1 vertical

SE and all edge pixels could be counted using a 3×3 solid SE.

While the above description is phrased in terms of subdividing the image and then processing each tile, the steps of Fig. 2B could be carried out on the entire image, and the result of XOR step 42 could be subdivided and pixels counted in each tile.

Fig. 3 is an expanded flow diagram illustrating the steps within step 12 (determining screen size and angle). Determination of the screen characteristics requires a set of spatial filters that are sensitive to halftone periodicity. These filters, to be described in detail below, are hit-miss SE's, each of which specifies a pattern of both ON and OFF pixels that is to be matched through erosion to the image. A set of N such filters that cover the frequency-angle space of expected printed halftones are selected (step 50) and the tile having the most pixel transitions as determined at step 27 is copied N times (step 52). Each copy is eroded with a respective filter (steps 55(1)...55(N)). The number of ON pixels remaining after erosion is counted for each filter (steps 57(1)...57(N)). The filter that yields the maximum number of ON pixels is selected (step 60), and provides the screen characteristics. As a further check to the actual existence of halftones, the maximum number can be compared to a threshold, and processing stopped if it does not exceed the threshold (step 62).

Figs. 4A-D, 5A-C, 6A-D, 7A-B, and 8A-D show the hit-miss SE's that are used as narrow bandpass filters to determine the halftone characteristics. Each filter consists of five ON pixels (hits) and four OFF pixels (misses) in a cruciform pattern with the ON pixels at the center and corners of a square and the OFF pixels approximately midway between the center and the corners. These filters have been found empirically to have a spatial period bandwidth of about 0.1 times the period of the repeating pattern and an angle bandwidth of about 15° .

Fig. 9A is a flow diagram illustrating the steps within step 15 (creation of separation mask). The image is copied (step 65), with one copy being used to create a halftone seed (step 67) and the other being used to create a clipping mask (step 70). The halftone seed must contain pixels only in the halftone regions and must contain at least one pixel in every halftone region. The clipping mask must densely cover all ON pixels in the halftone regions. While it may also cover non-halftone regions, no part of the clipping mask that covers a non-halftone region should touch a part of the clipping mask that covers a halftone region.

The seed and the clipping mask are then subjected to a series of operations that causes the seed to grow into the clipping mask (step 72). As will be discussed below, the seed and the clipping mask are generated at reduced scale. Accordingly, if it is desired to make direct use of the mask, the image resulting at step 72 is expanded to original scale (step 75).

Fig. 9B is an expanded flow diagram illustrating the steps within step 67 (creation of halftone seed). The image is first eroded with the best filter chosen at step 60 (step 80) and the result is twice reduced with $SCALE = 2$ and $LEVEL = 1$ (steps 82 and 83). Regions of text and line graphics, to the extent they survived the filtering, are blob-like at this point, but generally have small spatial extent. The image is then subjected to an open operation with a solid 3×3 SE (step 85), which removes any surviving pixels from the non-halftone regions. The result is a seed consisting of ON pixels only in regions corresponding to halftone regions in the original image.

Fig. 9C is an expanded flow diagram illustrating the steps within step 70 (creation of the clipping mask). The original image is subjected to a set of operations that eliminate OFF pixels that are near ON pixels. While text and lines are thickened, they tend to retain their general character. However, as the small dots in the textured regions expand, they coalesce to form large masses and thereby solidify the formerly textured area.

More particularly, the original image is twice reduced with $SCALE = 2$ and $LEVEL = 1$ (steps 90 and 92), which tends to solidify the halftone regions. The result is then optionally subjected to a close operation with a solid 2×2 SE (step 95). The result is a clipping mask which covers all halftone regions (and possibly other regions as well).

Fig. 9D is an expanded flow diagram illustrating the steps within step 72 (growing seed and clipping to mask). The seed, on the first iteration and thereafter an iterated image, is copied (step 100), and one copy reserved for future use. The other copy is dilated with a solid 3×3 SE (step 102), and the result ANDed with the clipping mask (step 105). The dilation tends to fill the seed while the AND operation ensures the seed does not grow beyond the clipping mask. The result of the AND operation is XORed with the reserved copy, to determine whether it has changed since the last iteration (step 107). This is done with an XOR operation, with a result of $XOR = 0$ (no ON pixels) implying that no change has occurred since the last iteration and the seed has filled the mask.

If the images differ, the current, partially filled seed may be copied at step 100 and the sequence continued until the XOR leaves no ON pixels. Alternatively, it may be desirable to stop the process after some number of iterations, say five, even if the image has not stopped changing. This could be the case where the filling activity represents a leak into a non-halftone region of the clipping mask that touches a halftone region inadvertently.

Fig. 10 is an expanded flow diagram illustrating the steps within step 20 (image separation). With the mask thus produced, segmentation is easily accomplished. The original binary image is copied (step 130) and ANDed with the mask (step 135) to produce the

"halftone separation" (i.e. the textured part of the image). This is copied (step 137) and the result is subjected to an exclusive OR (step 140) with a copy of the original binary image to produce the non-textured part of the image (text and line graphics). It is noted that the copying steps are not necessary if the result of the logical operation at any given stage is stored in a new array in memory rather than overwriting the operand.

Discussion of Illustrative Alternatives

Fig. 11 is a flow diagram illustrating an alternative to expansion of the solid portions to form the mask. Instead of expanding the solid regions to form the mask, it is possible to extract the coordinates of the solid regions (step 160) and scale the coordinates to full size (step 161). This provides a compact representation that allows convenient storage of the mask information.

The locations of the corners of each solid rectangular portion can be extracted by eroding the copies of the mask with a series of four SE's 162(ULC), 162(URC), 162(LLC), and 162(LRC). SE 162(ULC) picks out the upper left corner when it is used to erode a rectangle. The other SE's pick out the other corners.

For rectangular regions, it is useful, but not necessary, to extract the corner coordinates in a known order; the coordinates alone dictate the association of corners with rectangular regions. However, for non-rectangular regions, where there are more than four corners for each connected region of the mask (and there may be holes), the corners must be maintained in the order they are encountered by tracing around the periphery. Although possible, this may be sufficiently complicated that it is preferable to use the mask itself.

Fig. 12 is flow diagram illustrating an alternative to the technique of Fig. 9B for creating the halftone seed. The original binary image is twice reduced with SCALE = 2 and LEVEL = 1 (steps 170 and 172). The result is subjected to an optional close operation (step 175) with a solid 2×2 SE. The result is then twice further reduced with SCALE = 2 and LEVEL = 4 (steps 177 and 180). The result is then subjected to an OPEN operation with a solid 3×3 SE. It is noted that the seed produced by this sequence of operations is reduced by a factor of 16. Accordingly, the seed must be expanded by a factor of 4 to match the scale of the clipping mask, or the clipping mask must be reduced by a factor of 4.

Fig. 13A is an expanded flow diagram illustrating an alternative to the thresholded reductions with LEVEL = 1 used to create the seed and clipping mask. More particularly, the image is dilated with a solid 2×2 SE (step 190) and the resultant image sub-sampled by choosing one pixel in each 2×2 square to form a reduced image (step 192). The sub-sampling may be accomplished on a row basis by discarding every

other line, and on the column basis by use of a lookup table in a manner similar to that described below in connection with performing fast thresholded reductions.

Fig. 13B is an expanded flow diagram illustrating an alternative to the thresholded reductions with LEVEL = 4 used to create the seed. More particularly, the image is eroded with a solid 2×2 SE (step 195) and the resultant image sub-sampled by choosing one pixel in each 2×2 square to form a reduced image (step 197).

In principle one could also use a series of close operations in an attempt to solidify a finely textured region. However, the use of one or more thresholded reductions (or subsampling) has at least two advantages. First, because the texture scale is not known a priori, it cannot be determined how large an SE to use in the close operation. An SE that is too small to bridge adjacent parts in the textured region would not change the image, and the close operation would fail. Thus, while the use of the close would be locally all or nothing, the use of a reduction with LEVEL = 1 (or a dilation and subsampling) invariably results in a darkening of the texture. Second, the use of reductions before close allows the operation to be carried out at a reduced scale. Comparable operations at full scale are much slower computationally than those at a reduced scale (roughly by the third power of the linear scale factor). Therefore, all subsequent operations at reduced scale are much faster.

The selection of a particular size for the SE is empirically based, taking into consideration such factors as scanner resolution. However, the decision tends to be rather straightforward, entailing a minimum amount of experimentation.

Although the thresholded reductions are considered advantageous for creating the clipping mask, the same result is likely to be achieved by doing a dilation with a solid 8×8 SE followed by 8×8 subsampling. However, the computation is likely to take much longer.

Graphical Illustration

Fig. 14 shows a representative binary image scanned at 300 pixels/inch having text and halftone regions. Certain other views of the images at intermediate stages of processing are magnified 2x relative to Fig. 14. References to magnification in the following discussion are relative to Fig. 14 and are not to be confused with the expansion and reduction scale operations.

Fig. 15A is a plot of the filter response (number of ON pixels) resulting from erosion steps 55(1)...55(N) and counting steps 57(1)...57(N) illustrated in Fig. 3. As can be seen from the plot of Fig. 15A, filter number 15 has the greatest response. This filter actually corresponds to the filter of Fig. 8A, which has a spatial

period of eight pixels and an angle of 0°.

Fig. 15B shows the result (at 2x magnification) of erosion of the original image with the filter providing the maximum response. Figs. 15C and 15D show the results (also at 2x magnification) of the two subsequent reductions with SCALE = 2 and LEVEL = 1. The result is the seed.

Figs. 16A-B show the results (at 2x magnification) of the two reductions with SCALE = 2 and LEVEL = 1 carried out on the original image. As can be seen, the effect of the reductions is to cause the small texture elements within the halftone region to coalesce and solidify and to cause the text to thicken. The result is the clipping mask.

Fig. 17A shows the result (at 2x magnification) of filling the seed to the clipping mask. Fig. 17B shows the mask expanded to full size, that is by a linear factor of four corresponding to the two reductions each at SCALE = 2. This view is shown at the same scale as the original image.

Figs. 18A and 18B show the halftone and text separations. The halftone separation of Fig. 18A results from the logical AND of the mask and the original image while the text separation of Fig. 18B results from the exclusive OR of the halftone separation and the original image, as shown in Fig. 10.

Fast Thresholded Reduction (and Expansion) of Images

One requirement of efficient segmentation, is that thresholded reduction must be done quickly. Suppose it is desired to reduce an image by a factor of two in the vertical direction. One way to do this is to use a raster operation (bitblt - bit block transfer) to logically combine the odd and even rows, creating a single row of the reduced image for each pair of rows in the original. The same procedure can then be applied to the columns of the vertically squashed image, giving an image reduced by a factor of two in both directions.

The result, however, depends on the logical operations of the horizontal and vertical raster operations. Obtaining a result with LEVEL = 1 or 4 is straightforward. If an OR is used for both raster operation orientations, the result is an ON pixel if any of the four pixels within the corresponding 2 x 2 square of the original were ON. This is simply a reduction with LEVEL = 1. Likewise, if an AND is used for both raster operation orientations, the result is a reduction with LEVEL = 4, where all four pixels must be ON.

A somewhat different approach is used to obtain a reduction with LEVEL = 2 or 3. Let the result of doing a horizontal OR followed by a vertical AND be a reduced image R1, and let the result from doing a horizontal AND followed by a vertical OR be a reduced image R2. A reduction with LEVEL = 2 is obtained by ORing R1 with R2, and a reduction with LEVEL = 3 is obtained by ANDing R1 with R2.

The procedure may not be computationally efficient if implemented as described above. On some computers, such as Sun workstations, raster operations are done in software. The image is stored as a block of sequential data, starting with the first row of the image, moving left-to-right, then the second row, etc. Consequently, the raster operations between rows are fast, because 16 or 32 bits in two words can be combined in one operation. But to perform a raster operation between two columns, the corresponding bits must be found, two bits at a time (one from each column), before the logical operations can be done. It turns out that the time, per pixel, to do the vertical raster operations is at least 25 times greater than the horizontal ones. In fact, when the algorithm is implemented entirely with raster operations, over 90 percent of the time is devoted to the vertical operations.

Fortunately, there is a simple and very fast way to implement the logical operations between columns. Rather than use column raster operations, take 16 sequential bits, corresponding to 16 columns in one row. These 16 bits can be accessed as a short integer. These 16 bits are used as an index into a 2¹⁶-entry array (i.e. a lookup table) of 8-bit objects. The 8-bit contents of the array give the result of ORing the first bit of the index with the second, the third bit with the fourth... and on to the 15th bit with the 16th. Actually, two arrays are needed, one for ORing the 8 sets of adjacent columns, and one for ANDing the columns. It should be understood that the numerical example is just that, an example. It is also possible to implement this as a 2⁸-entry array of 4-bit objects, or any one of a number of other ways.

The use of lookup tables to implement column logical operations is about as fast, per pixel, as Sun's row raster operations. A 1000 x 1000 pixel image can be reduced on a Sun 3/260, with either LEVEL = 1 or 4, to a 500 x 500 pixel image in 0.1 seconds. On a Sun 4/330, the operation takes about 0.04 second.

Special Hardware Configuration

As discussed above, 2 x 2 reductions require a first logical operation between rows followed by a second, possibly different, logical operation between columns. Moreover, some threshold levels require two intermediate reduced images which are then combined. The table lookup technique for column operations can become cumbersome if it is desired to have a very wide pixelword. Either the table becomes enormous or one needs special techniques of looking up parts of the wide pixelword in multiple parallel tables. The latter, while clearly superior, does require some way to use portions of the data word as memory addresses, which may not otherwise be necessary.

Fig. 19 is a logic schematic of specialized hardware for performing a logical operation between

vertically adjacent 2Q-bit pixelwords and a pairwise bit reduction of the resulting 2Q-bit pixelword (bits 0 through 2Q-1). Although the drawing shows a 16-pixel word, the benefits of this hardware would become manifest for much longer pixelwords where the lookup table technique has become cumbersome. A 512-bit pixelword is contemplated, since a line of image would represent only a few pixelwords.

The reduction of the two pixelwords occurs in two stages, designated 200 and 202. In the first stage, a vertically adjacent pair of pixelwords is read from a first memory 203, and the desired first logical operation is carried out between them. The desired second logical operation is then carried out between the resulting pixelword and a version of the pixelword that is shifted by one bit. This provides a processed pixelword having the bits of interest (valid bits) in every other bit position. In the second stage, the valid bits in the processed pixelword are extracted and compressed, and the result stored in a second memory 204. Memory 203 is preferably organized with a word size corresponding to the pixelword size. Memory 204 may be organized the same way.

The preferred implementation for stage 200 is an array of bit-slice processors, such as the IDT 49C402 processor, available from Integrated Device Technology. This specific processor is a 16-bit wide device, each containing 64 shiftable registers. Thirty-two such devices would be suitable for a 512-bit pixelword. For simplification, a 16-bit system with four registers 205, 206, 207, and 208 is shown. Among the processor's operations are those that logically combine the contents of first and second registers, and store the result in the first. The processor has a data port 215, which is coupled to a data bus 217.

Second stage 202 includes first and second latched transceivers 220 and 222, each half as wide as the pixelword. Each transceiver has two ports, designated 220a and 220b for transceiver 220 and 222a and 222b for transceiver 222. Each transceiver is half as wide as the pixelword. Ports 220a and 222a are each coupled to the odd bits of data bus 217, which correspond to the bits of interest. Port 220b is coupled to bits 0 through (Q-1) of the data bus, while port 222b is coupled to bits Q through (2Q-1). The bus lines are pulled up by resistors 125 so that undriven lines are pulled high.

Consider the case of a 2×2 reduction with LEVEL = 2. The sequence of operations requires that (a) a vertically adjacent pair of pixelwords be ANDed to form a single 2Q-bit pixelword, adjacent pairs of bits be ORed to form a Q-bit pixelword, and the result be stored; (b) the vertically adjacent pair of pixelwords be ORed, adjacent bits of the resultant 2Q-bit pixelword be ANDed, and the resultant Q-bit pixelword be stored; and (c) the two Q-bit pixelwords be ORed.

To effect this, a pair of vertically adjacent pixelwords are read from first memory 203 onto data bus

217 and into registers 205 and 206. Registers 205 and 206 are ANDed and the result stored in registers 207 and 208. The content of register 208 is shifted one bit to the right, registers 207 and 208 are ORed, and the result is stored in register 208. Registers 205 and 206 are ORed, and the result stored in registers 206 and 207. The content of register 207 is right shifted by one bit, registers 206 and 207 are ANDed, and the result stored in register 207.

At this point, register 207 contains the result of ORing the two pixelwords and ANDing pairs of adjacent bits, while register 208 contains the result of ANDing the pixelwords and ORing pairs of adjacent bits. However, registers 207 and 208 contain the valid bits in the odd bit positions 1, 3,... (2Q-1). For a reduction with LEVEL = 2, registers 207 and 208 are ORed and the result is made available at processor data port 215 which is coupled to data bus 217.

The odd bits of the data bus are latched into transceiver 220 through port 220a, resulting in a Q-bit pixelword with the valid bits in adjacent positions. Although this Q-bit entity could be read back onto the bus and transferred to memory 204, it is preferable to use both latches. Thus, two new pixelwords (horizontally adjacent to the first two) are processed at stage 200 as described above, the result is made available at processor data port 215, and is latched into transceiver 222 through port 222a. The contents of the two transceivers are then read out through ports 220b and 222b onto data bus 217 in order to provide a 2Q-bit pixelword that represents the reduction of four 2Q-bit pixelwords. The result is transferred to second memory 204. This overall sequence continues until all the pixelwords in the pair of rows has been processed. Once the pair of rows has been processed, subsequent pairs are similarly processed.

As mentioned above each bit-slice processor has 64 registers. Accordingly, since memory accesses are more efficient in a block mode, faster operation is likely to result if 8 pairs of pixelwords are read from memory 203 in a block, processed as discussed above, stored in the processor's registers, and written to memory 204 in a block.

Image enlargement is similar, but the steps are executed in the reverse order. First, the processor reads a pixelword and sends the left half through port 220b of transceiver 220. This is read onto the bus through port 220a. Only every other pixel in the resulting word on the bus will initially be valid, so the processor will need to validate all the pixels using a sequence of shifts and logic operations. Since resistors 125 pull up all the bus lines that are not driven, each undriven line, all the even bits in this case, will be 1's. This expanded pixelword, which alternates 1's with valid data, is read into two registers, the content of one register is shifted one place, and the registers are logically ANDed. Everywhere there was a 0 in an odd bit, there will be 00 in an even/odd pair. None of

the other bits will be affected. This pixelword is then written to two vertically adjacent words in the expanded image. This process is repeated for the right half of the pixelword using the transceiver 222. The processor expands the entire row one pixelword at a time and the entire image one row at a time.

The techniques described above with reference to the drawings provides for efficient and effective separation of textured regions from non-textured regions in an image.

They are robust and computationally efficient techniques utilizing transformations, such as thresholded reduction operations and morphological operations, for determining the existence of halftone or stippled regions (referred to collectively as halftone regions) in a binary image, characterizing such regions as to screen size and angle, and creating a representation of the halftone regions, such as by creating a separation mask covering the halftone regions or extracting the coordinates of the halftone region boundaries. The techniques described with reference to the drawings are computationally simple since the operations are essentially local. They are also fast. This is in part because the operations entail few computations per pixel, and in part because many of the operations are carried out on reduced images, which contain fewer pixels than the original image, so there are fewer pixels to process.

Various modifications of the described embodiments are possible. For example, all the above operations could be done on an image, that is first reduced. Thus, much of the image computation would occur at a further reduced scale. Moreover, while the specific original image illustrated above was derived from a document scanned at 300 pixels/inch, the disclosure applies to documents scanned at any other resolution.

Claims

1. A method for constructing a separation mask for halftone regions in an original binary image, comprising the steps of :
 constructing a seed ;
 constructing a clipping mask ; and
 growing the seed and clipping to the clipping mask (Fig. 9A).
2. A method as claimed in claim 1, and further comprising the step of determining the halftone characteristics.
3. A method as claimed in claim 1 or claim 2, and further comprising the steps of :
 providing a set of spatial filters ; and
 selecting the spatial filter that best matches the halftone characteristics (Fig. 3).
4. A method as claimed in claim 3, wherein said selecting step comprises the steps of :
 eroding at least a portion of the image with each of the filters ;
 counting the ON pixels for each erosion ; and
 choosing the filter that gives the largest number.
5. A method as claimed in claim 3 or claim 4, wherein said step of constructing a seed comprises :
 eroding the original binary image with the selected spatial filter ;
 subjecting the image, thus filtered, to at least one thresholded reduction ; and
 subjecting the resulting image to an open operation (Fig. 9B).
6. A method as claimed in any one of claims 3 to 5, wherein each of the filters is a hit-miss structuring element having a pattern of five ON pixels located at the center and corners of a square and four OFF pixels located generally midway between the center and corners of the square (Figs. 4 to 8).
7. A method as claimed in claim 1, wherein said step of constructing a seed comprises the steps of :
 subjecting the original binary image to at least one thresholded reduction at a low threshold value ;
 subjecting the resulting image to at least one thresholded reduction at a high threshold level ; and
 subjecting the resulting image to an open operation (Fig. 12).
8. A method as claimed in any one of the preceding claims, wherein said step of constructing a clipping mask comprises subjecting the image to a set of operations that eliminates OFF pixels that are near ON pixels (Fig. 9C).
9. A method as claimed in any one of the preceding claims, wherein said steps of growing and clipping comprise :
 on a first pass copying the seed to define a current version ;
 dilating the current version ;
 ANDing the clipping mask and the current version to define an iterated version ;
 comparing the copy of the current version with the iterated version ; and
 (a) if they are equal exiting, or
 (b) if they are unequal, copying the iterated version to redefine the current version and repeating the above steps, including this one, starting at said dilating step (Fig. 9D).
10. A method as claimed in any one of the preceding

claims, and further comprising the steps, performed before said seed construction step, of :

determining whether halftone regions exist in the original binary ; and
 exiting if and only if halftone regions do not exist. 5

11. A method as claimed in claim 10, wherein said step of determining whether halftone regions exist comprises the steps of : 10

dividing the image into subregions ;
 determining the approximate number of pixel transitions in each subregion ;
 picking the subregion with the greatest number of pixel transitions ; and 15
 comparing that number to a threshold, above which a subregion is considered to be largely halftoned.

12. A method as claimed in claim 11, wherein said step of determining the approximate number of pixel transitions comprises the steps of : 20

eroding the subregion with a horizontal SE of adjacent ON pixels ;
 XORing the result with the subregion ; and 25
 counting ON pixels remaining (Fig. 2B).

13. A method as claimed in claim 4 and claim 11, in which the portion of the image that is eroded with each of the filters is the subregion with the greatest number of pixel transitions. 30

14. A method of separating halftone regions of a binary image, comprising the steps of constructing a separation mask by a method as claimed in any one of the preceding claims, and logically combining the separation mask and the original binary image to extract image separations (Fig. 10). 35

40

45

50

55

11

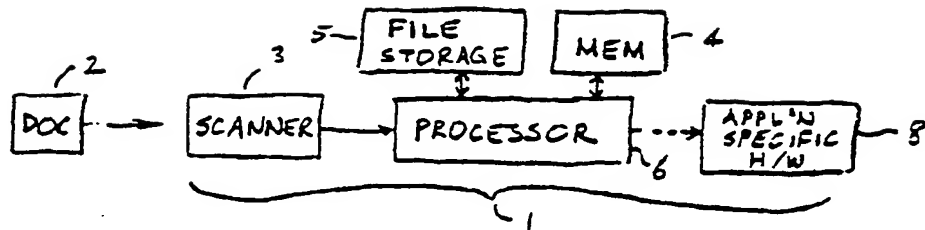


FIG. 1A

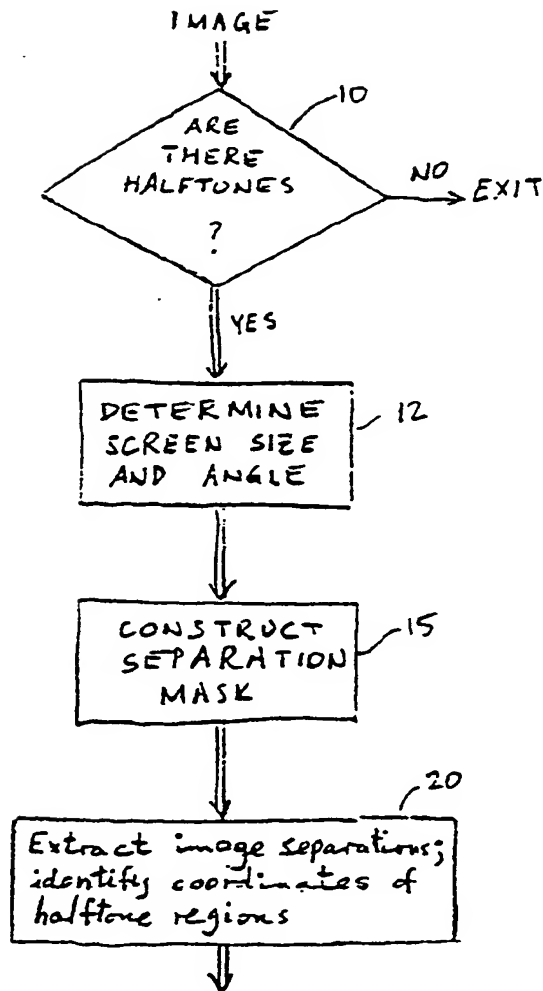


FIG. 1B

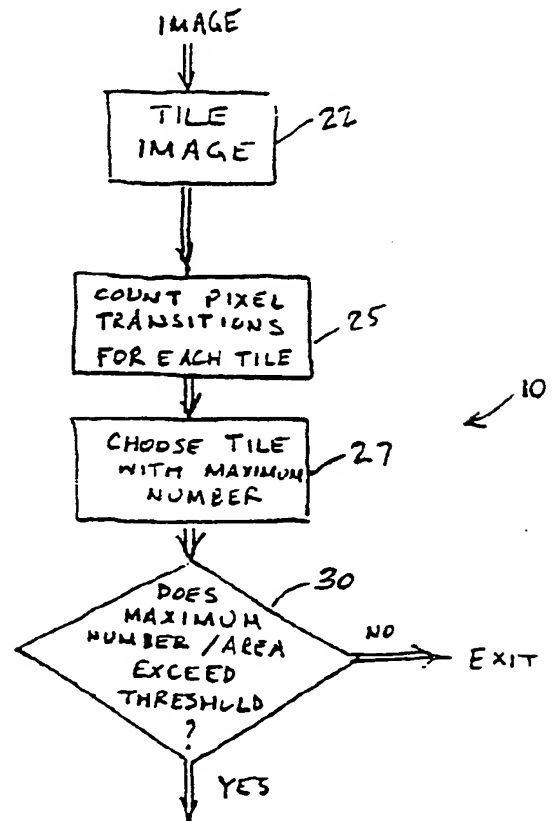


FIG. 2A

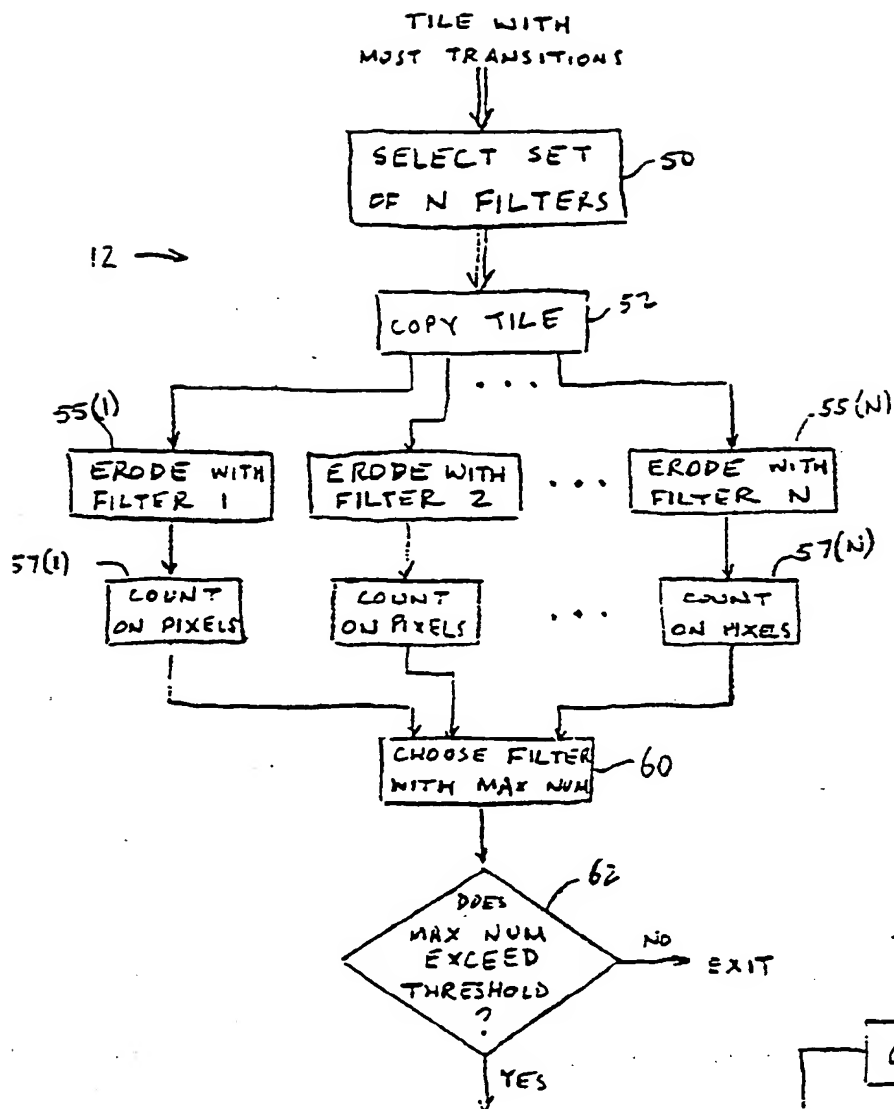


FIG. 3

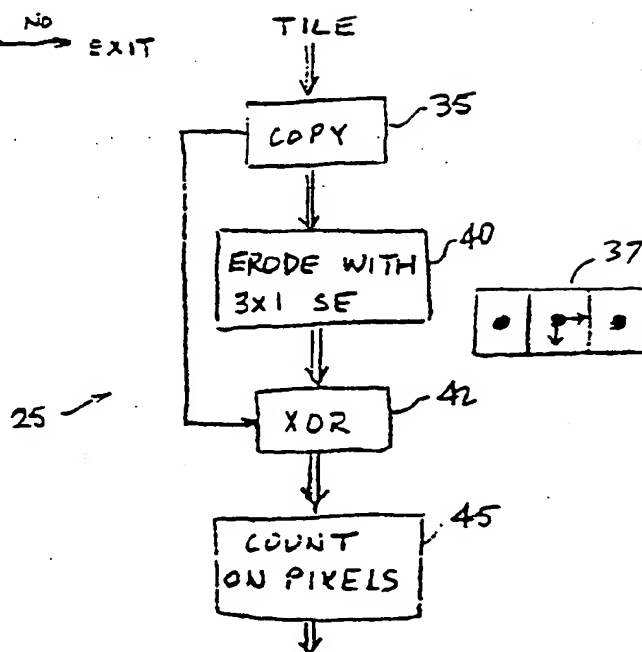
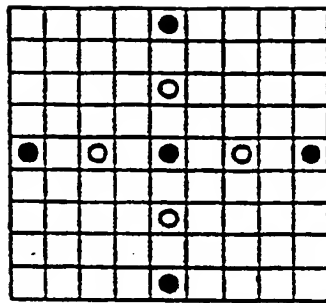
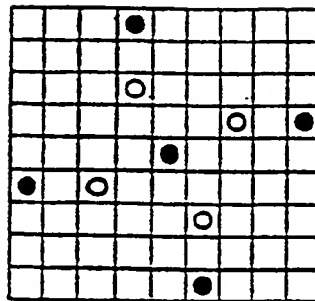


FIG. 2B



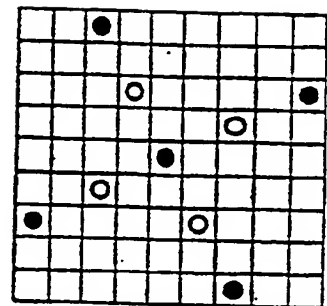
period = 4, angle = 0 degrees

FIG. 4A



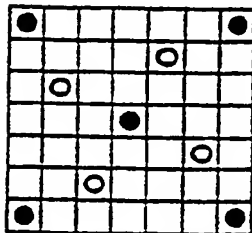
period = 4, angle = 14 degrees

FIG. 4B



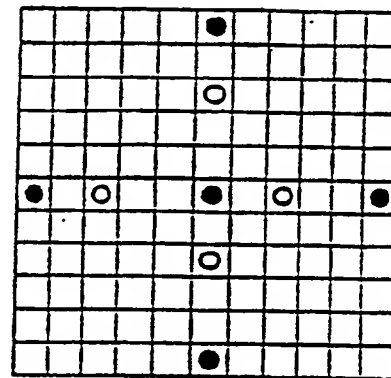
period = 4, angle = 27 degrees

FIG. 4C



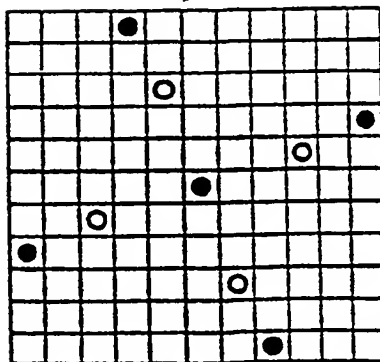
period = 4, angle = 45 degrees

FIG. 4D



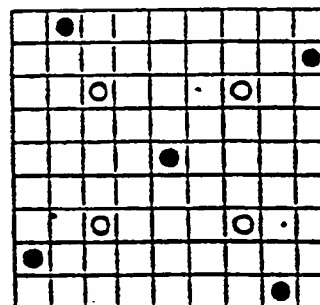
period = 5, angle = 0 degrees

FIG. 5A



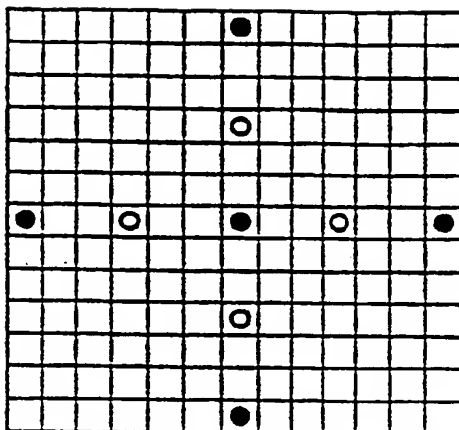
period = 5, angle = 22 degrees

FIG. 5B



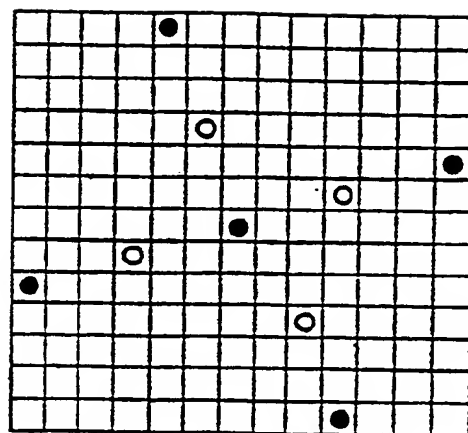
period = 5, angle = 37 degrees

FIG. 5C



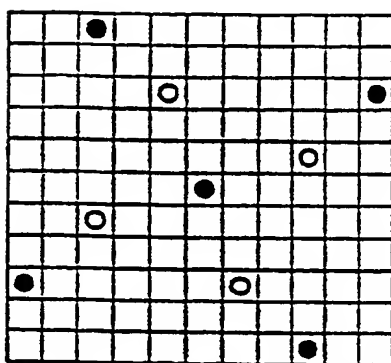
period = 4, angle = 9 degrees

FIG. 6A



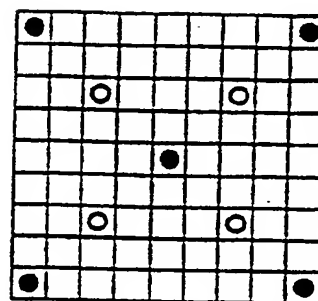
period = 6, angle = 18 degrees

FIG. 6B



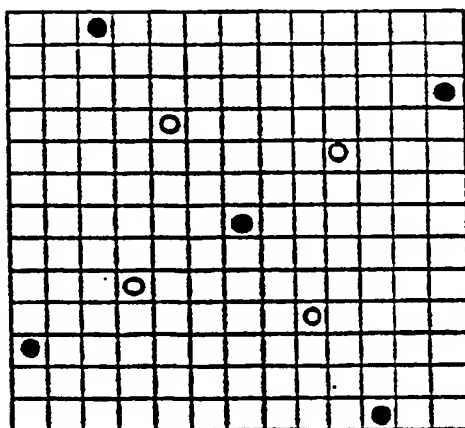
period = 6, angle = 31 degrees

FIG. 6C



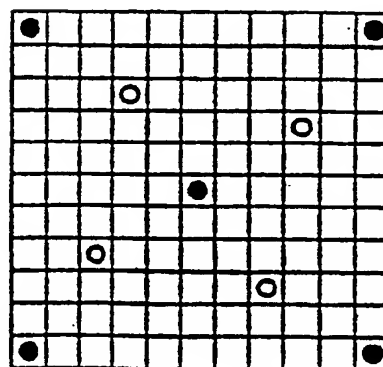
period = 6, angle = 45 degrees

FIG. 6D



period = 7, angle = 34 degrees

FIG. 7A



period = 7, angle = 45 degrees

FIG. 7B

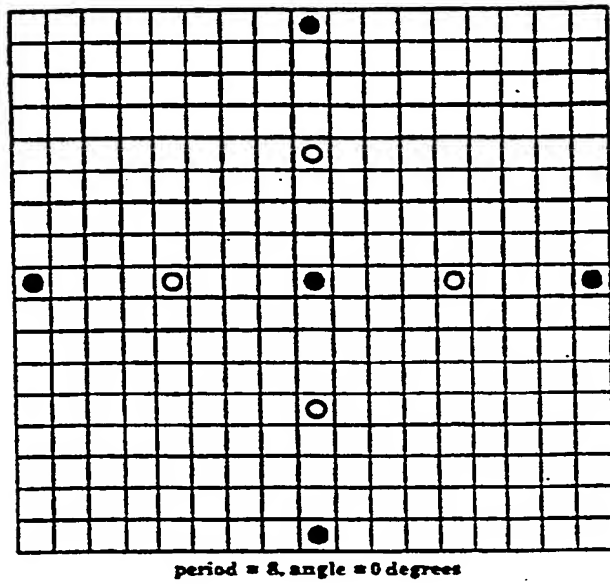


FIG. 8A

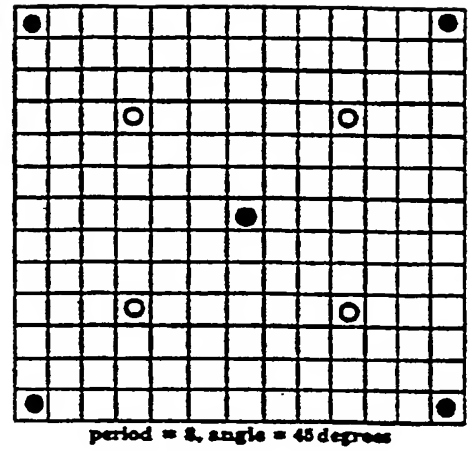


FIG. 8D

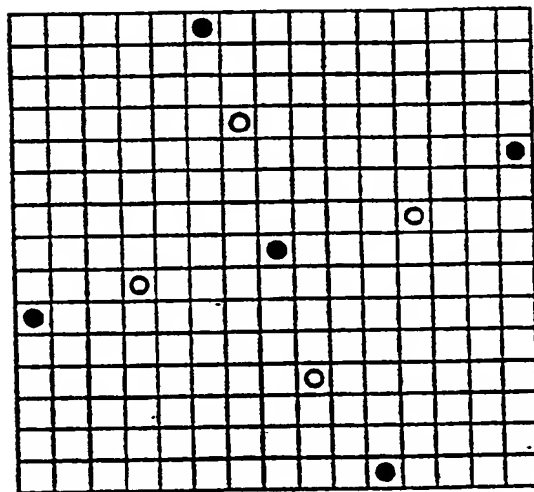


FIG. 8B

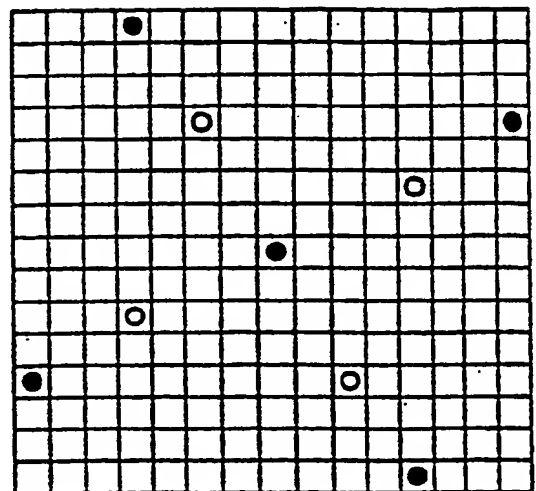
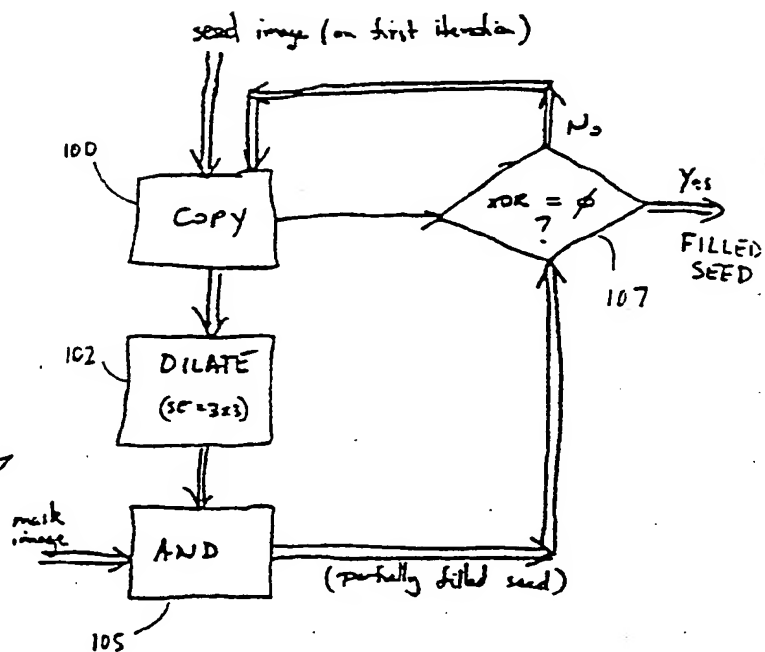
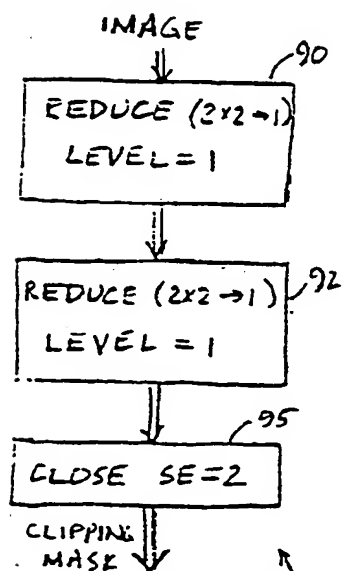
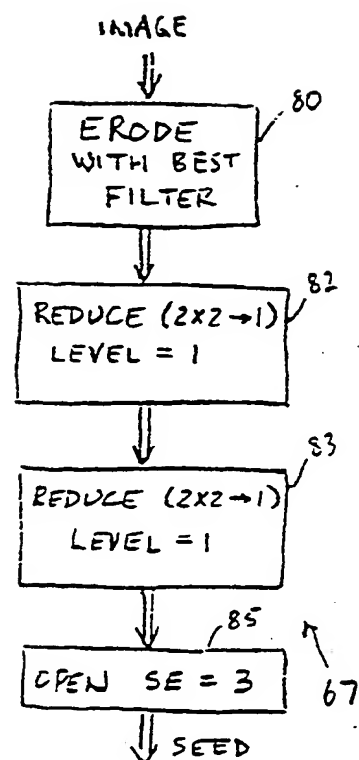
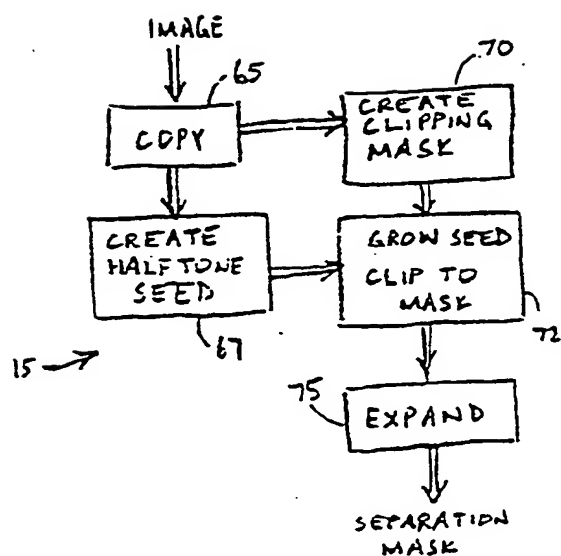


FIG. 8C



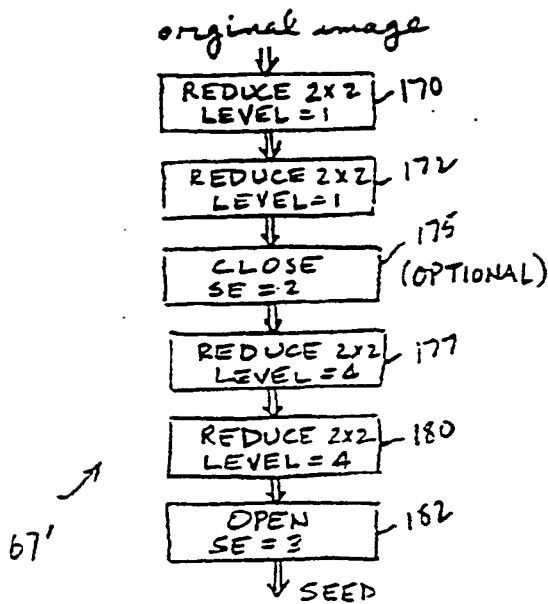


FIG. 12

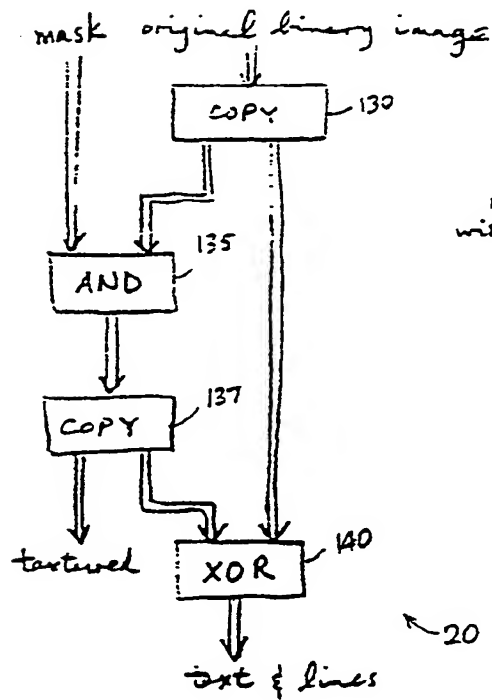


FIG. 10

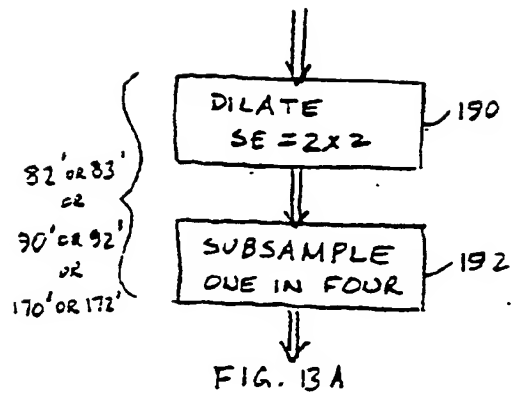


FIG. 13A

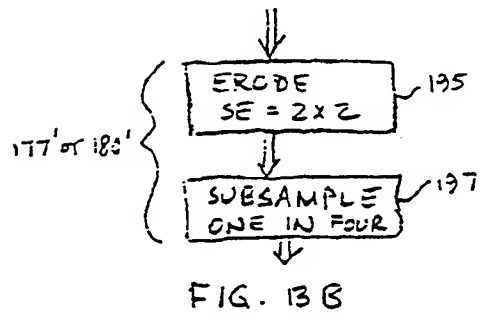


FIG. 13B

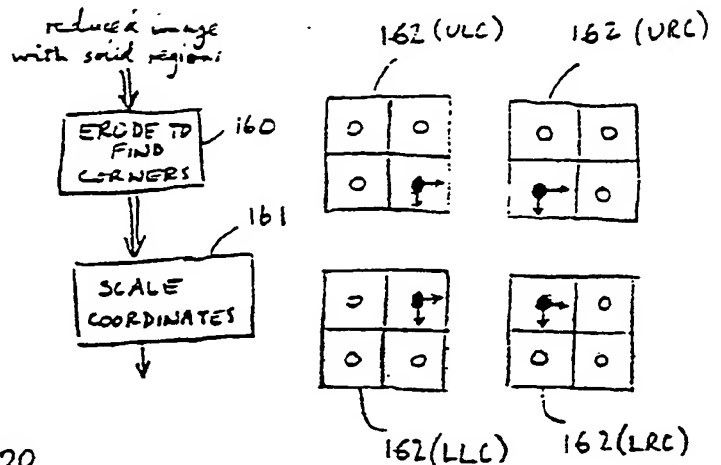
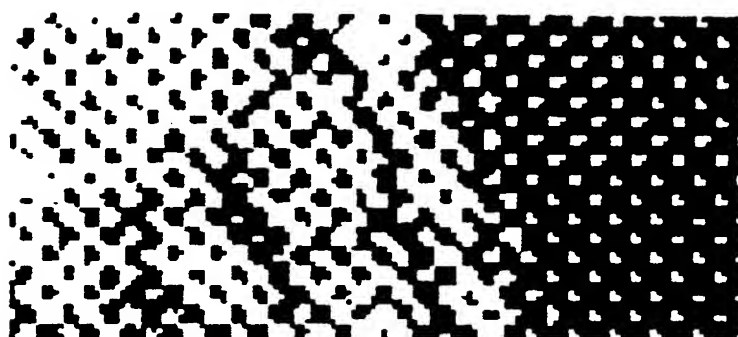


FIG. 11



3CDEI

cdefgh

FIG. 14

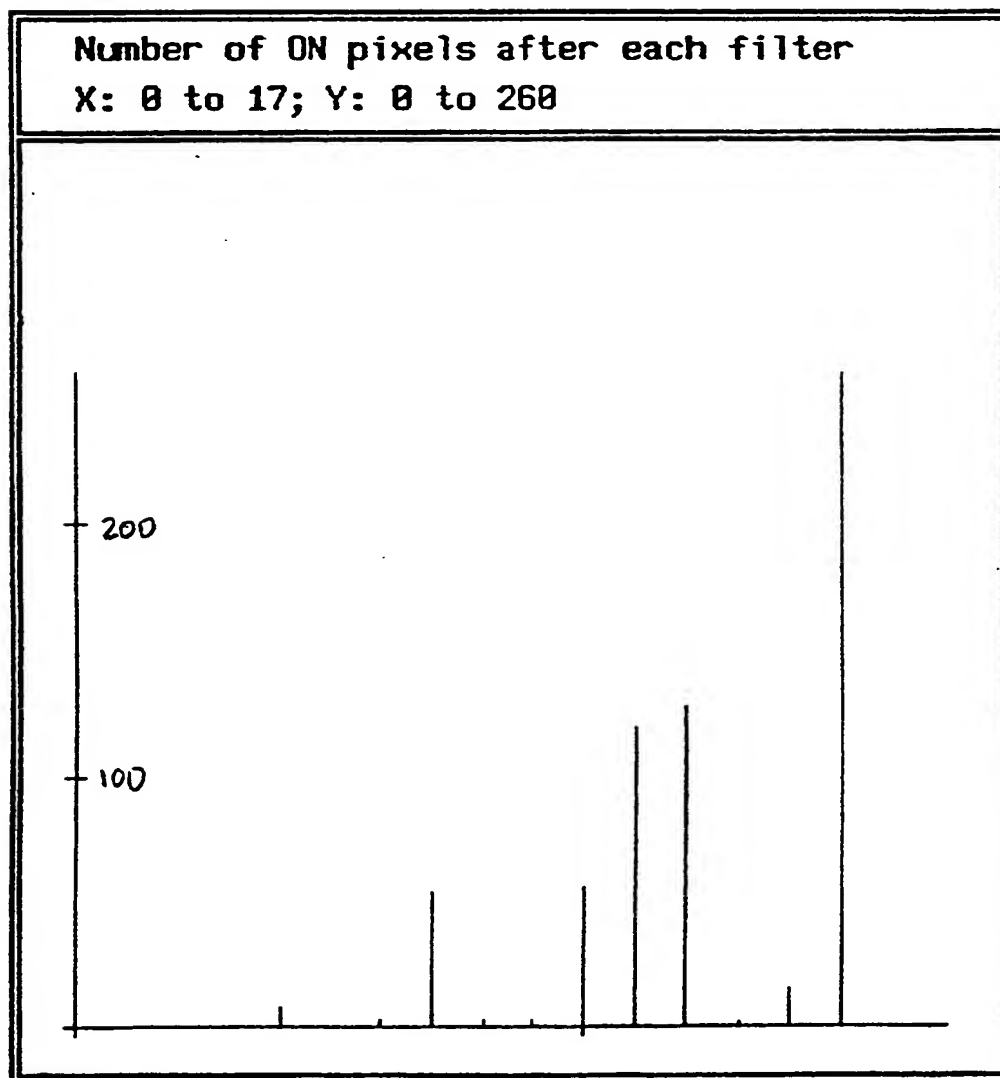


FIG. 15A

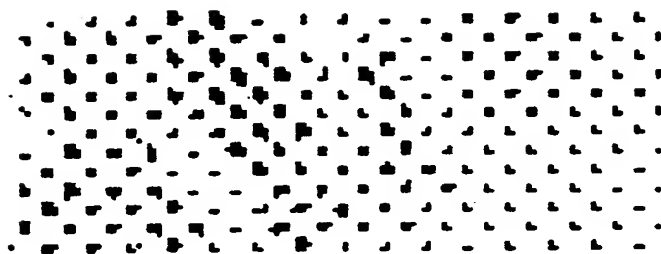


FIG. 15B

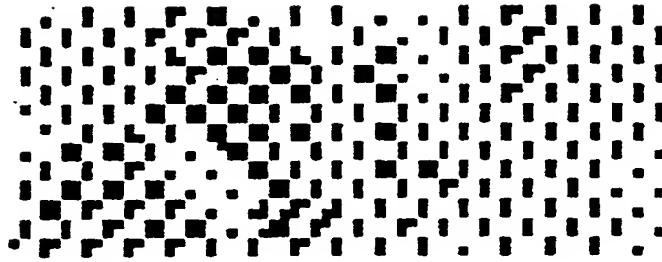


FIG. 15C

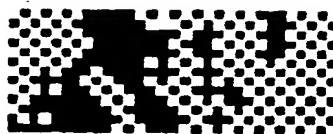
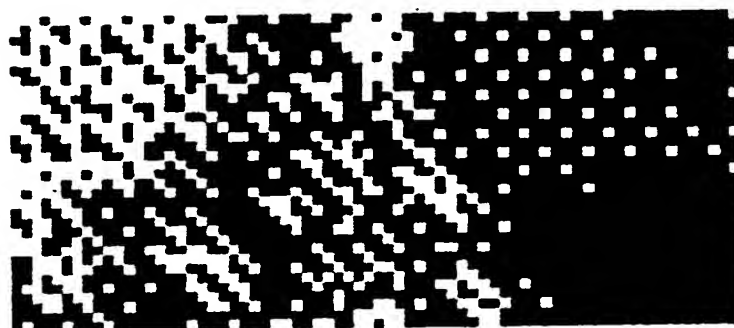


FIG 15D



3CDEI
cdefgh

FIG. 16A



FIG. 17A



CODE

cdygh

FIG. 16B

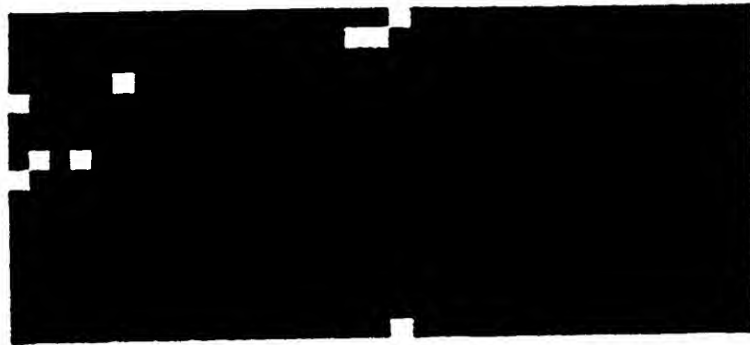


FIG. 17B

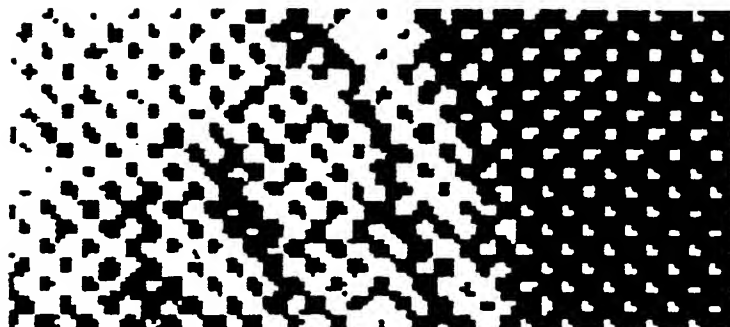


FIG. 18 A

3 C D E
c d e f g h

FIG. 18 B

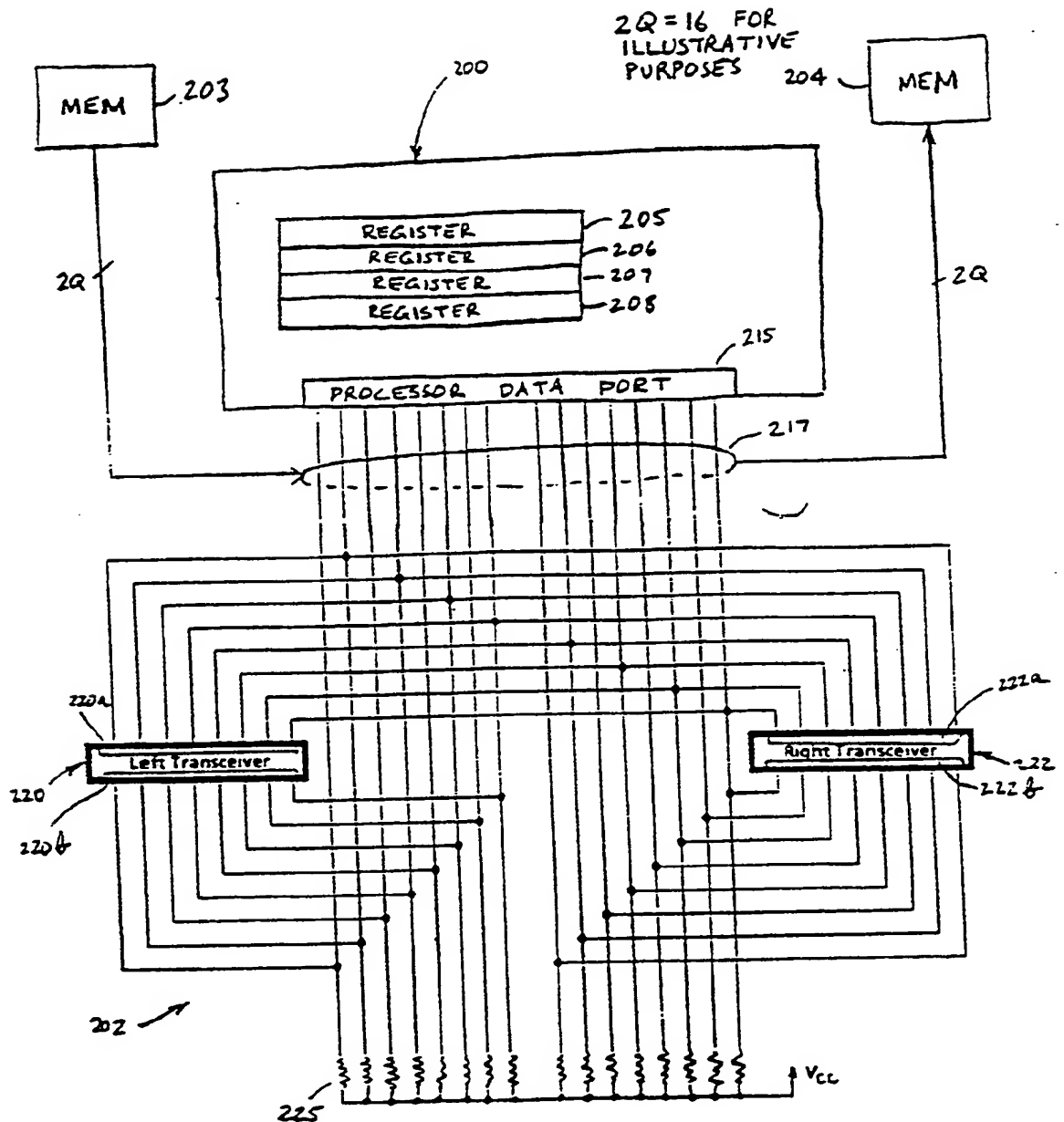


FIG. 19



12 **EUROPEAN PATENT APPLICATION**

21 Application number : **90313311.4**

51 Int. Cl.⁵ : **H04N 1/40, G06F 15/68**

22 Date of filing : **07.12.90**

30 Priority : **08.12.89 US 448193**

43 Date of publication of application :
12.06.91 Bulletin 91/24

84 Designated Contracting States :
DE FR GB

88 Date of deferred publication of search report :
25.03.92 Bulletin 92/13

71 Applicant : **XEROX CORPORATION**
Xerox Square - 020
Rochester New York 14644 (US)

72 Inventor : **Bloomberg, Dan S.**
1013 Paradise Way
Palo Alto, California, 94306 (US)
Inventor : **Crocca, William T.**
114 Heather Way
Rochester, New York, 14625 (US)

74 Representative : **Hill, Cecilia Ann et al**
Rank Xerox Patent Department Albion House,
55 New Oxford Street
London WC1A 1BS (GB)

54 Identification, characterization, and segmentation of halftone or striped regions of binary images.

57 A method for creating a mask for separating halftone regions in a binary image from other regions comprises : constructing a seed image that includes pixels only in halftone regions and at least one pixel in every halftone region (67) ; constructing a clipping mask that covers in a connected manner all ON pixels in halftone regions (70) ; and filling the seed while clipping to the mask (72). Thresholded reductions and morphological operations are preferred.

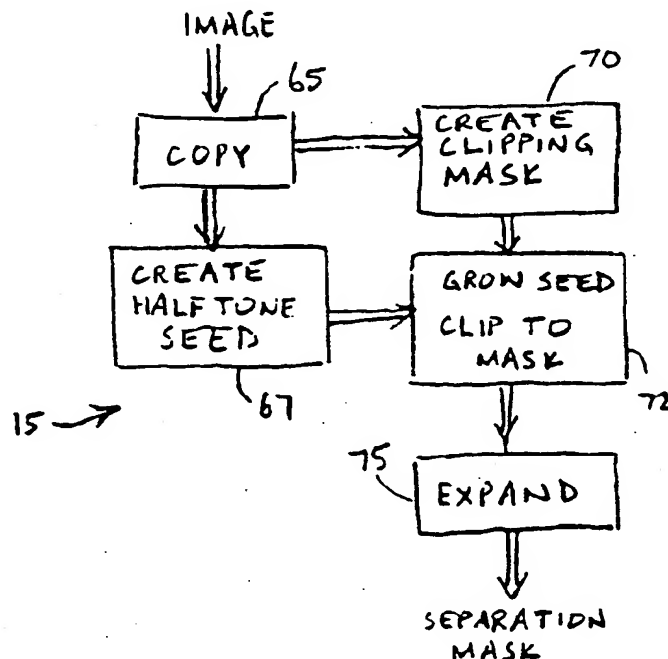


FIG. 9A



European Patent
Office

EUROPEAN SEARCH REPORT

Application Number

EP 90 31 3311
Page 1

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int. Cl.5)
A	XEROX DISCLOSURE JOURNAL. vol. 5, no. 6, December 1980, STAMFORD, CONN, US page 643; R.S. TOMORY: 'Halftone detection using transition counting' * complete document *	1, 4, 11, 12	H04N1/40 G06F15/68
A	IBM TECHNICAL DISCLOSURE BULLETIN. vol. 27, no. 7B, December 1984, NEW YORK, US pages 4279 - 4282; S.J. FOX ET AL.: 'High frequency information detector/discriminator' * complete document *	2, 3, 6	
A	RADIO MENTOR ELECTRONIC. vol. 40, no. 3, 1974, MUNCHEN, DE pages 105 - 108; G. KAMIN: 'Der Geometrie-Computer; Ein neues Konzept der Video-Bildanalyse' * complete document *	4, 5, 8, 9, 12, 13	
A	IEEE TRANSACTIONS ON PATTERN ANALYSIS AND MACHINE INTELLIGENCE. vol. 7, no. 6, November 1985, NEW YORK, US pages 638 - 651; S. SUZUKI ET AL.: 'New fusion operations for digitized binary images and their applications' * sections III, VI * * figures 4, 5, 10, 12 *	1, 3, 6, 10, 14	TECHNICAL FIELDS SEARCHED (Int. Cl.5) H04N G06F
A	12th CANADIAN SYMPOSIUM ON REMOTE SENSING. vol. 3, 10-14 July 1989, IEEE, NEW YORK, US pages 1220-1223; E.R. DOUGHERTY ET AL.: 'Image segmentation by local morphological granulometries' * complete document *	1, 10, 14	
--- -/--			
The present search report has been drawn up for all claims			
Place of search THE HAGUE		Date of completion of the search 29 JANUARY 1992	Examiner HOEKSTRA A.
CATEGORY OF CITED DOCUMENTS X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document		T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons A : member of the same patent family, corresponding document	

EPO FORM 1503 (2.12.1991) (P0401)



European Patent
Office

EUROPEAN SEARCH REPORT

Application Number

EP 90 31 3311

Page 2

DOCUMENTS CONSIDERED TO BE RELEVANT			
Category	Citation of document with indication, where appropriate, of relevant passages	Relevant to claim	CLASSIFICATION OF THE APPLICATION (Int. CL.5)
A	12th CANADIAN SYMPOSIUM ON REMOTE SENSING, vol. 3, 10-14 July 1989, IEEE, NEW YORK, US pages 1209-1213; I. TANNOUS ET AL.: 'SAR images segmentation using region growing approach' * complete document *	1, 14	
			TECHNICAL FIELDS SEARCHED (Int. CL.5)
The present search report has been drawn up for all claims			
Place of search THE HAGUE		Date of completion of the search 29 JANUARY 1992	Examiner HOEKSTRA A.
<p>CATEGORY OF CITED DOCUMENTS</p> <p>X : particularly relevant if taken alone Y : particularly relevant if combined with another document of the same category A : technological background O : non-written disclosure P : intermediate document</p> <p>T : theory or principle underlying the invention E : earlier patent document, but published on, or after the filing date D : document cited in the application L : document cited for other reasons & : member of the same patent family, corresponding document</p>			

EPO FORM 1503 01.82 (P0401)

THIS PAGE BLANK (USPTO)